

AD-A068 206

TEXAS A AND M UNIV COLLEGE STATION INST OF STATISTICS F/6 12/1
ONESAM, A COMPUTER PROGRAM FOR NONPARAMETRIC DATA ANALYSIS AND --ETC(U)
MAR 79 E PARZEN, J M WHITE DAA629-78-6-0180

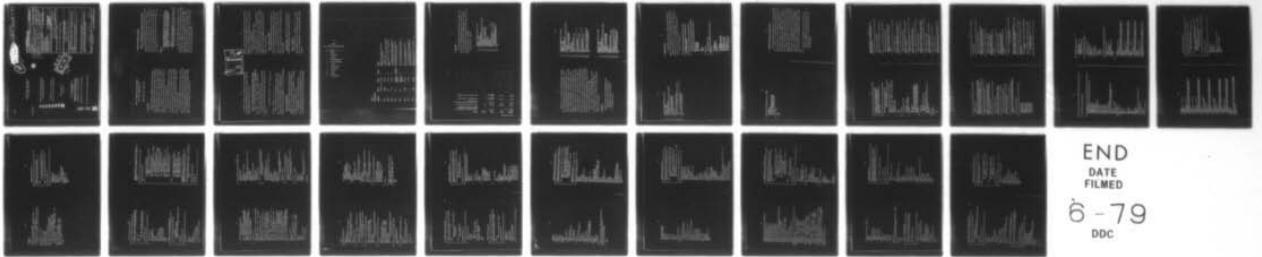
UNCLASSIFIED

TR-A-7

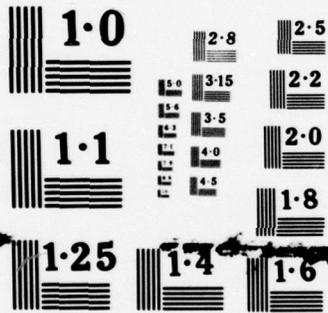
ARO-16228.7-M

NL

1 OF 1
ADA
068206



END
DATE
FILMED
6-79
DDC



NATIONAL BUREAU OF STANDARDS
MICROCOPY RESOLUTION TEST CHART

TEXAS A&M UNIVERSITY
COLLEGE STATION, TEXAS 77843

INSTITUTE OF STATISTICS
FORM 711, 5-63-111

ONESAM, A COMPUTER PROGRAM FOR NONPARAMETRIC
DATA ANALYSIS AND GOODNESS OF FIT

by Emanuel Parzen and J. Michael White
Institute of Statistics, Texas A&M University

Technical Report No. A-7
March 1979

Texas A & M Research Foundation
Project No. 3564

"Maximum Robust Likelihood Estimation and
Non-parametric Statistical Data Modeling"
Sponsored by the U.S. Army Research Office

Professor Emanuel Parzen, Principal Investigator

Approved for public release; distribution unlimited.

AD A068206

DDC FILE COPY

DDC
REF ID: A71473
MAY 3 1979
REGISTRY

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Technical Report No. A-7	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ONESAM, A Computer Program for Nonparametric Data Analysis and Goodness of Fit	5. AUTHOR(s) Emanuel Parzen and J. Michael White	6. PERFORMING ORG. REPORT NUMBER 9 Technical rept.
7. AUTHORING ORG. Texas A&M University Institute of Statistics College Station, TX 77843	8. CONTRACT OR GRANT NUMBER(s) 15 DAAG29-78-C-0190	9. PERFORMING ORG. REPORT NUMBER
10. CONTROLLING OFFICE NAME AND ADDRESS Army Research Office Research Triangle Park, NC 27709	11. MONITORING AGENCY NAME & ADDRESS (if different from Contracting Office) 11 March 1979	12. SECURITY CLASS. (of this report) Unclassified
13. DISTRIBUTION STATEMENT (of this report) Approved for public release; distribution unlimited.	14. DISTRIBUTION STATEMENT (of this report) 14 TR-A-7	15. SECURITY CLASSIFICATION (of this report) Unclassified
16. SUPPLEMENTARY NOTES The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.	17. KEY WORDS (Continue on reverse side if necessary and identify by block number) Statistical data modeling, statistical computing, goodness of fit tests	18. AVIS RACT (Continue on reverse side if necessary and identify by block number) This report is a Users' Guide to ONESAM, a mainline Fortran computer program that implements the one-sample non-parametric data analysis and Goodness of Fit procedures introduced by Parzen (1979).

DD FORM 1473
1 JAN 75
EDITION OF 1 NOV 68 IS OBSOLETE
S/N 0102-10-011-6401

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

347380 Jul

12
BEST AVAILABLE COPY
18
19
ARL 16228.7-M

ONESAM, A COMPUTER PROGRAM FOR NONPARAMETRIC

DATA ANALYSIS AND GOODNESS OF FIT

by

Emanuel Parzen and J. Michael White

1. Introduction

ONESAM (one sample) is a FORTRAN mainline program that implements the one sample non-parametric data analysis and Goodness of Fit techniques proposed by Parzen (1979). There are 5 stages to a quantile function and density-quantile function analysis of a single data sample. The program: Stage 1. Accepts grouped or individual (ungrouped) data, and forms non-parametric raw estimates of the quantile function; Stage 2. Forms a raw density-quantile function and sample spacings.

If the user specifies a null hypothesis probability law, the program checks the goodness of fit of the data to that probability law, and forms non-parametric smooth estimates of the density-quantile function in Stages 3 and 4. Stage 5 forms $\hat{Q}(u)$, a smooth estimator of the quantile function when a parametric model is assumed.

Section 2 describes the stages of analysis in more detail. Section 3 details the procedure to use ONESAM including the options available to users. In Section 4, several examples of the JCL used to run the jobs are given. Section 5 discusses Quantile-Box Plots, and Section 6 introduces a program QANMAT for parametric estimation of location and scale.

2. Comments on the Stages of Analysis

Stage 1: User inputs data, outputs sample quantile function.

A. If the input is ungrouped data, the user must supply: the tape number (specified by DATA), upon which the data resides (if other than unit 5 which is the tape the data is usually on when read in from cards); the format of the data; and the data itself. The output of stage 1 consists of: various descriptive statistics; the order statistics by quartiles; and the empirical quantile function, $\hat{Q}(u)$, defined by $\hat{Q}(u) = X_{[nu]}$ for $u = \frac{j-0.5}{n}$, $j = 1, \dots, n$.

Options:

1. Data can be standardized to take values in the unit (0, 1) interval by specifying IOPT2 = 1. The output (which includes descriptive statistics, order statistics by quartiles, and \hat{Q}) is based on the standardized data. When this option is used, future stages of analysis will use standardized data for analysis.
2. Compute $\hat{Q}(u)$ using linear interpolation, for $u = jh$, $j = 1, \dots, NQ$ where $NQ = [1/h]$ by specifying FINC = h. When this option is used, future stages will use \hat{Q} computed at points $h, 2h, \dots, 1-h$.
3. If the user inputs grouped data by specifying IOPT1 = 1, the user must also supply: the lower limit of the first interval (EMIN); the interval width (XINT); and, if desired, the total number of observations (NOBS). The user must input the numbers of observations in each interval in Format (20P4.0) and supply a value, h (FINC), to be used in computing \hat{Q} . As with the raw data, the user must supply the number of the unit (DATA) that the data is to be read from (if other than unit 5).

ACCESSION for		White Section		NHS	
ANNOUNCED		<input type="checkbox"/> Bill Section		ODC	
JUSTIFICATION		<input type="checkbox"/>		BY	
DISTRIBUTION/AVAILABILITY CODES		SPECIAL		A 23	

Also available in Stage 1 is the Quantile Box Plot procedure which is not part of the ONESAM package but is easy to access by a short user-supplied FORTRAN program described in Section 5.

Stage 2: Raw density-quantile function $\hat{f}(u)$ and sample spacings.

In Stage 2, the function $\hat{f}(u)$, defined by $\hat{f}(u) = \frac{1}{h} [\bar{Q}((j+1)h) - \bar{Q}(jh)]$ for $u = jh, j = 1, \dots, m$, is computed. Also computed in Stage 2 is $\hat{q}(u) = 1/\hat{f}(u)$. Plots of these functions are given.

Note that if \bar{Q} is not computed via linear interpolation from grouped data, but rather is computed from ungrouped data, the definition of $\hat{q}(u)$ is such that

$$\hat{q}(u) = n(\bar{Q}((j + .5)/n) - \bar{Q}((j - .5)/n)) = n(X_{(j+1)} - X_{j:n})$$

for $\frac{j-0.5}{n} \leq u < \frac{j+0.5}{n}, j = 1, 2, \dots, n-1$.

Stages 1 and 2 constitute the non-parametric raw analysis of the data. If the user supplies a null-hypothesis probability law, one can proceed to Stages 3 and 4 which constitute a non-parametric smooth analysis of the data and also a goodness of fit test for the null hypothesis.

Stage 3: Raw $\hat{D}(u)$ and $\hat{f}(v)$.

The user must specify a density-quantile function $f_0 q_0$ for the null hypothesis $H_0: Q = \mu + \sigma Q_0$. This is done using indicator variables. If one wishes to test against density "j", then set $bQR(j) = 1$. The possible densities one can use are supplied in Table I.

The program computes: (1) the raw transformation density,

$\hat{d}(u) = f_0 q_0(u) \hat{q}(u) / \hat{\sigma}_0$, also called the weighted spacings; and (2) the raw transformation distribution function, $\hat{D}(u)$ called the cumulative weighted

spacings. (Tests for H_0 could be obtained by using the facts that under H_0 the \hat{d} 's should be distributed as Uniform (0, 1), and $\sqrt{n}(\hat{D}(u) - u)$ is asymptotically a Brownian Bridge stochastic process). $\hat{\sigma}_0$ is an estimate of σ_0 defined by $\hat{\sigma}_0 = \int_0^1 f_0 q_0(u) \hat{q}(u) du$.

The program then computes the Fourier transform

$\hat{\phi}(v) = \int_0^1 \hat{d}(u) \exp(2\pi i v u) du$ for $v = 0, 1, \dots, m$ by specifying ORDER = m (the program chooses order 5 if no other order is specified). By comparing the square modulus of the phi's, $|\hat{\phi}(v)|^2$, which are plotted, to a suitable threshold value (for example, 2*NOBS), one could use these values as evidence to accept or reject H_0 .

Stage 4: Smooth $\hat{D}(u)$ and $\hat{f}(u)$.

Autoregressive smoothed density estimators $\hat{f}(u)$ and smoothed transformation distribution functions $\hat{D}(u)$ are computed for successive orders 1, 2, ..., m. Plots of \hat{D} vs \hat{D} are given for all orders to help select the order that gives the best fit to the data. Also given is the CAT criterion. If CAT selects order 0, one uses this as evidence to not reject the null hypothesis.

Alternative analyses: Stages 3 and 4 can be repeated for several possible hypothesized densities simply by setting any other $bQR(j) = 1$.

Stage 5: Smooth $\hat{Q}(u)$.

The program computes $\hat{Q}(u) = \hat{\mu} + \hat{\sigma} Q_0(u), u = jh, j = 1, \dots, m$ for several estimates $\hat{\mu}, \hat{\sigma}$ of μ and σ . This procedure is available but is not part of the ONESAM package; it is a separate program QANHEAT described in Section 6.

TABLE I

List of Density-Quantile Functions

j	$f_0 Q_0$
1	Normal
2	Exponential
3	Logistic
4	Double Exponential
5	Uniform Reciprocal
6	Cauchy
7	Extreme Value
8	Log Normal
9	Parato
10	Weibull
11	Half Logistic

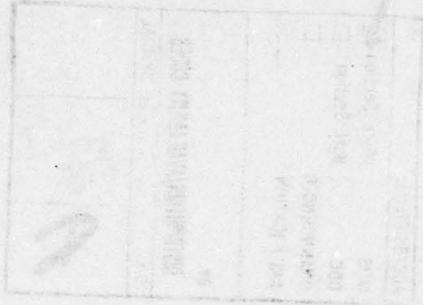


TABLE II

Parameter List

Parameter	Value	Status	Default	Remarks
N =	integer ≤ 510	required	0	Sample size for raw data; the number of intervals for grouped data
DATA =	integer	optional	5	1 = data on unit 1 (if DATA = 2, card 14 must change accordingly)
TRANS =	real	optional	1.0	Value of exponent for power transformation if desired
IOP1 =	0, 1, 2	optional	0	1 for grouped data with equal interval widths 2 for grouped data in unequal intervals
XMIN =	real	optional required if IOP1 = 1	first order statistic	Natural minimum of data. For grouped data, it is lower limit of first interval
FINC =	real	optional required if IOP1 = 1	0.0	Value of h for computing equally spaced Q_i . If used, FINC usually equals .01, .025, or .05
NOBS =	integer	optional	N	Number of observations for grouped data. It is used in computing CAT
IOP2 =	0, 1	optional	0	1 if user desires data to be standardized to (0, 1) before analysis. Sometimes useful for comparing several batches
ORDER =	integer	optional	5	Maximum order to be used for autoregressive estimator usually ≤ 10

3. Procedures

This section details the procedure one should use to complete a data analysis using the ONESAM package and gives the JCL necessary to run a job. The most critical feature of the ONESAM package is to understand completely the parameter list PARMs for this contains almost all the information and options used in the analysis.

A typical JCL deck of cards is:

```
1. //XXXX JOB (XXXX,XXX,XXX,XXX,XXX,XXX),XXXXX
2. //PASSWD XXXX
3. //JOBPARM R=192,K=0
4. //EXEC FORTX08,FXRGM=192K
5. //FORT.SYSIN DD DS=U.L.MJ.EMP.QUANT(ONSPL1),DISP=SHR
6. //GO.SYSL IB DD
7. // DD DS=U.L.MJ.EMP.COSUB1,DISP=SHR
8. // DD DS=U.L.MJ.EL.M.CTSDD,DISP=SHR
9. //GO.SYSIN DD *
10. SA MARCH=III,SEND
11. DATA SET NAME
12. SPANMS N=III,SEND
13. FORTM OF DATA EG (16F5.1)
14. //GO.FTOF001 DD DS=U.L.MJ.EMP.QUANT(ONSPL1),DISP=SHR
15. //DISP=SHR,LABEL=(,IM)
16. //GO.FTOF001 DD UNIT=SYSDA,SPACE=(TRK,50),
17. //DCB=(RECFM=F3,LRECL=80,BLKSIZ=3120)
18. /*END
```

TABLE II (continued)

Parameter	Value	Status	Default	Remarks
DQM(J) = J = 1, ... 12	0, 1	optional	0	1 If user wishes to test against "J"th hypothesized density. See table in Section 3 for allowable densities
WIDTH =	Integer	optional	80	Number of columns in printer plots
PLOT =	0, 1	optional	0	1 If CAT comp plots desired (unavailable at present time)
BATAF =	real > 0	required if	D(9) = 1	Shape parameter for Pareto density
BETAJ =	real > 0	required if	D(10) = 1	Shape parameter for Weibull density

Cards 10, 12, 13, 14, and 15 require further explanation. Cards 10 and 12 utilize the name list feature of the FORTRAN language. Card 10 tells the program how many batches are being analyzed in this run. If only one batch is being analyzed, the card should be punched: `B4A SEND` (Note: `b` indicates a blank space.) If more than one batch is being analyzed, the user would repeat cards 11 through 15 `REATCH` times. The user must put a blank in column 1 of card 10 and an ampersand in column 2. Card 12 is the parameter card. The user must put a blank in column 1 of card 12 and an `s` in column 2. Columns 3-7 contain the word `PARMS` followed by a blank in column 8. The rest of the card contains the desired parameters selected from the parameter list in Table II. Some of the parameters are required, some are optional as explained below. They can be listed in any order. The parameter list must end with `,SEND`.

Card 13 is required if the user inputs ungrouped data. Omit card 13 if using grouped data. Cards 14 and 15 are used if the data is located on an external field. If the data is on cards, replace cards 14 and 15 with the data cards.

4. Sample JCL

Sample JCL are given for the following types of jobs:

- A. Ungrouped data on cards, with three batches;
- B. Ungrouped data on external file (Wylbur);
- C. Grouped data on cards;

```

A
1. //WHITE JOB (E463,0029,530,005,MH), 'BUFSM0'
2. //PASSWD= XXXX
3. //JOBPARM N=172, K=0
4. // EXEC FORTXCO, FORTGM=172K
5. // FORT. SYS IN DD DSN=UPL.MI.DMP.QUANT (ON SPL 1), DISP=SHR
6. //60.SYSLIB DD
7. // DD DSN=UPL.MI.DMP.CO.SUB1.DISP=SHR
8. // DD DSN=UPL.MI.DMP.CT.SBD.DISP=SHR
9. //60.SYS IN DD *
10. SA NBATCH=3,SEND
11. TIPPETT'S WARP BREAKS AL
12. SPARMS N=9,DOH(1)=1,DOH(2)=1,END
13. (10FS.0)
14. 24. 30. 54. 25. 70. 52. 51. 26. 67.
15. TIPPETT'S WARP BREAKS AM
16. SPARMS N=9,DOH(1)=1,DOH(2)=1,SEND
17. (10FS.0)
18. 21. 29. 17. 12. 10. 35. 30. 36.
19. TIPPETT'S WARP BREAKS AM
20. SPARMS N=9,DOH(1)=1,DOH(2)=1,SEND
21. (10FS.0)
22. 36. 21. 24. 18. 10. 43. 28. 15. 26.
23. //60.F109F001 DD UNIT=SYSDA,SPACE=(TRK,50),
24. // DCB=(RECFM=FB,URECL=60,RLSEIZE=3120)
25. //END

```

```

B
1. //WHITE JOB (E463,0029,530,005,MH), 'BUFSM0'
2. //PASSWD= XXXX
3. //JOBPARM N=172, K=0
4. // EXEC FORTXCO, FORTGM=172K
5. // FORT. SYS IN DD DSN=UPL.MI.DMP.QUANT(ONESPL1), DISP=SHR
6. //60.SYSLIB DD
7. // DD DSN=UPL.MI.DMP.CBSUP1.DISP=SHR
8. // DD DSN=UPL.MI.DMP.CT.SBD.DISP=SHR
9. //60.SYS IN DD *
10. SA SEND
11. BIFFALD SMOULFALL DATA (1910-1972)
12. SPARMS N=63,DOH(1)=1,INT=1,FILE=05,SEND
13. (SF19.0)
14. //60.F101F001 DD DSN=UPL.MI.DMP.QUANT(BFSM0),
15. // DISP=SHR, LABEL=(,1,1)
16. //60.F109F001 DD UNIT=SYSDA,SPACE=(TRK,50),
17. // DCB=(RECFM=FB,URECL=60,RLSEIZE=3120)
18. //END

```

```

C
1. //DATE JOB (683,0028,S30,005,M1), 'BP DATA'
2. //PASSED XDDX
3. //CEPANI R=172, E=0
4. // EXEC FOR TICS, PREGN=172K
5. //FCST SYS IN DD DSN=U.L.M. EMP QUANT (ON SPL1), DISP=SHR
6. //SD.SLSLIB DD
7. // DD DSN=U.L.M. EMP.COSURS, DISP=SHR
8. // DD DSN=U.L.M. BLN.CTSBB, DISP=SHR
9. //DD.SYS IN DD
10. //
11. //JOB P-RESUME DATA--NON-USERS--#SE 17-24
12. //SEGS N=15, B(1)=1, DATA=5, IOPT1=1, XMIN=BS, XMIT=5, FINC=.05,
13. //
14. //
15. //FC.FIC9F001 DD UNIT=SYS3A, SPACE=(TRK,50),
16. // DCB=(RECFB=F8, LRECL=80, BLKSIZE=3120)
17. //*END

```

5. Quantile Box-Plot

A simple FORTRAN driver program to access the subroutines QDIAC and QBOX to perform the Quantile Box plot procedures and compute the associated diagnostics is supplied below. If the user desires a Calcplot of the Quantile Box Plot, he can use the subroutine QBXPLT. Be cautioned that the data vectors and vector of "q-values" must be dimensioned at least 2 * N + 1 where N is the sample size.

```

C-----
C
C THIS PROGRAM IS THE DRIVER FOR THE QUANTILE BOX PLOT
C ANALYSIS AND DIAGNOSTICS.
C-----
C DIMENSION DATA(200), X1(200), X2(200), I(200), IU(200), LABEL(20)
C
C DATA LABEL/40DUFF,4HALD,4MSHOW,4HIFALL,4N 191,4NO-72,
C 4100AN /
C DATA LAB1/4HURIB/
C DATA LAB2/4HSORT/
C DATA LAB3/4HLOG /
C
C N=43
C NI=2*N+1
C IOP1=0
C
C OBTAIN DATA
C
C READ(5,100) (DATA(I), I=1,N)
C
C COMPUTE PRELIMINARY TRANSFORMATIONS OF DATA FOR ANALYSIS
C IF DESIRED
C
C DO 10 J=1,N
C X1(J)=SORT(DATA(J))
C X2(J)=ALOG(DATA(J))
C 10 CONTINUE
C
C LABEL(8)=LAB1
C CALL QDIAC(DATA,N,IOP1,LABEL)
C CALL QBOX(DATA,Y,NI,LABEL)
C LABEL(8)=LAB2
C CALL QDIAC(X1,N,IOP1,LABEL)
C CALL QBOX(X1,Y,NI,LABEL)
C LABEL(8)=LAB3
C CALL QDIAC(X2,N,IOP1,LABEL)
C CALL QBOX(X2,Y,NI,LABEL)

```

```

C C OBTAIN CALCCOMP OF QUANTILE BOX PLOT, IF DESIRED
C
C LABEL(8)=4H
C LAB=10CRIGINAL
C CALL OSPLIT(DATA, NW, LABEL, 3, J, 2, LAB1)
C CALL OSPLIT(X1, NW1, LABEL, 3, J, 2, LAB2)
C CALL OSPLIT(X3, NW1, LABEL, 3, J, 2, LAB3)
C
C 100 FOR NW1 (GF 10.0)
C
C STOP
C END

```

6. Parametric Estimation of Location and Scale

One can obtain parametric estimates $\hat{\mu}$ and $\hat{\sigma}$ of μ and σ in the model $Q(u) = \mu + \sigma Q_0(u)$. For a complete discussion of the derivation of $\hat{\mu}$ and $\hat{\sigma}$ see Parzen (1978). At the present time parametric estimates are available only when we assume $Q_0 = \Phi^{-1}$. Estimates $\hat{\mu}(p)$ and $\hat{\sigma}(p)$ can be computed using the subroutine STEPL, MURAT, and SICHAT. The variance of each estimator is also computed. The estimate $\hat{Q}_p(u) = \hat{\mu}(p) + \hat{\sigma}(p)\Phi^{-1}(u)$ can be computed using the estimates $\hat{\mu}(p)$ and $\hat{\sigma}(p)$ together with subroutine MDRIS which will compute $\Phi^{-1}(u)$. A plot of $\hat{Q}(u)$ vs $\hat{Q}(u)$ would be useful to check the fit of the parametric model and can be obtained using subroutine JFPLOT. A sample FORTRAN program to produce the estimates and desired plots is given below. This sample program uses ungrouped data as input and would have to be modified somewhat to accommodate grouped data.

Note: This stage of the analysis is in the process of revision and expansion to accommodate the full range of Q_0 functions. Provisions to make this stage of analysis more accessible to users are being developed.

INDEX OF SUBROUTINES USED IN THE ONESAM PACKAGE. UTILITY
SUBROUTINES ARE LISTED IN A SEPARATE GROUPING FROM
COMPUTING SUBROUTINES.

FUNCTION AREST(X,L,OPTRNA,OPTDDE)

FUNCTION TO COMPUTE AUTOREGRESSIVE ESTIMATOR EVALUATED
AT X.

SUBROUTINE AUTORG(A,L,S,M,ALPHA,PHI, RGH)

COMPUTES THE COEFFICIENTS ALPHA(I) AND FKH OF THE
AUTOREGRESSIVE ESTIMATOR ACCORDING TO A RECURSIVE
ALGORITHM

SUBROUTINE DATSTBX(XO,N,ON)

SUBROUTINE TO STANDARDIZE VECTOR X BY (X(I)-XO)/(X(M)-XO)
AND RETURN THE STANDARDIZED DATA IN THE VECTOR ON

SUBROUTINE DESTAT(X,N,NAME, IUNIT, IHEAD)

SUBROUTINE TO PRINT ORDERED ARRAY BY QUANTILES AND COMPUTE
DESCRIPTIVE STATISTICS.

SUBROUTINE FORTIER(F,U,H,A,MA)

SUBROUTINE TO COMPUTE THE FOURIER TRANSFORM
PHI(V) OF A DENSITY DEFINED ON (0,1) FOR V=0,1,...,M

SUBROUTINE KSD(B,U,N,DM,UM,DP,UP)

SUBROUTINE TO COMPUTE KOLMOGOROV-SMIRNOFF STATISTIC FOR
THE DEVIATIONS D(U)-U. UPPER AND LOWER BOUNDS ARE COMPUTED

SUBROUTINE PARZ(RVAR,N,N,CAT,NORD)

SUBROUTINE TO DETERMINE THE ORDER OF AN AUTOREGRESSIVE
PROCESS BY PARZEN'S CAT CRITERIA

SUBROUTINE PRRTA(A,N,IUNIT)

SUBROUTINE TO COMPUTE AND PRINT THE SQUARE ROOTS OF THE
COMPLEX-VALUED FOURIER TRANSFORMS A(1),...,A(IN)

C THIS PROGRAM IS THE DRIVER TO OBTAIN ESTIMATES
C RCHAT(P) AND SIGMAT(P) FOR VARIOUS VALUES OF P.
C IT ALSO COMPUTES CHAT(U) AND OTILNE(B) FOR
C UNGROUPED DATA, AND PLOTS GRAY AND OTILNE ON THE
C SAME SET OF AXES
C *****

C DIMENSION DATA(200),U(200),CHAT(200),PHI(200),
C LAB1(20),LAB2(20),LAB3(20),
C P(4),XNU(4),UNU(4),UNU(4),XSIG(4),VSIG(4),USIG(4)
C *****

C DATA LAB1/4HBUFF,4HALD,4HSHOU,4HMLL,4H 191,4H0-72,
C 4H /
C DATA LAB2/4HPL0T,4H OF,4HCHAT,4H(0),4HWS 0,4HTILD,4H(0),
C 4H /
C DATA LAB3/4HP = ,19*4H /
C DATA P/0.,.05,.1,.25/
C *****

C IP=4
C IP=3
C *****

C OBTAIN DATA
C *****

C READS,100)DATA(1),I=1,N)
C *****

C DO 10 I=1,N
C U(I)=(FLOAT(I)-.5)/FLOAT(N)
C CALL NDARIS(U(I),PHI(I),IER)
C 10 CONTINUE
C *****

C WRITE(4,101)
C DO 20 IP=1,MP
C CALL STEP1(P(IP),UNU(IP),USIG(IP),UNU(IP),VSIG(IP))
C CALL RCHAT(P(IP),DATA,N,UNU(IP),UNU(IP),UNU(IP))
C CALL SIGMAT(P(IP),DATA,N,USIG(IP),VSIG(IP),XSIG(IP))
C WRITE(4,102)(P(IP),UNU(IP),UNU(IP),XSIG(IP),VSIG(IP))
C 20 CONTINUE
C *****

C DO 40 IP=1,MP
C DO 30 I=1,M
C CHAT(I)=XNU(IP)+XSIG(IP)*PHI(I)
C 30 CONTINUE
C CALL FCODEA(9,P(IP),B(F6.4),LAB3(I))
C CALL JPLOT(CHAT,U,N,BO,4H(0),4H U,LAB2,LAB3,4,1,1,1,
C 4HDATA,1,LAB1)
C 40 CONTINUE
C *****

C 100 FORMAT
C 101 FORMAT(//5X,
C 102 FORMAT(
C *****

C STOP
C END

SUBROUTINE QUICKIN(T)

QUICK SORT THIS ALGORITHM IS ALSO REFERRED TO AS A PARTITIONED EXCHANGE SORT. EXPECTED RUNTIME IS PROPORTIONAL TO N*LOG2(N) ALTHOUGH THE WORST CASE IS PROPORTIONAL TO N**2. REFERENCE: DONALD E. KNUTH - THE ART OF COMPUTER PROGRAMMING VOL 3.

SUBROUTINE CHLIX(X,N,I1,U,O,IOP1,NO,XH,IN,XINT)

SUBROUTINE TO COMPUTE Q(U) AT U = N,2N,3N,...,I-H WHERE H IS > OR = .002 BY LINEAR INTERPOLATION FROM ORDER STATISTICS IF DATA IS UNGROUPED (IOP1=0), OR FROM TALLIES IF DATA IS GROUPED (IOP1=1).

SUBROUTINE QTOFO(Q,U,NO,XS,F0)

SUBROUTINE TO COMPUTE Q(U) AND F(Q)=1./Q(U) FROM THE EMPIRICAL QUANTILE FN CAP Q(U) AND THE U VALUES.

SUBROUTINE USPACE(CWS,CXNS,NO,FO,F0,NO,U,SIG0)

SUBROUTINE TO COMPUTE Q(U), CUMULATIVE D'S, AND SIGMA FOR THE MODEL Q(U)=NO*SIGMA*Q(U)

SUBROUTINE ONESAM(X,N,X0,OM,XS,CXNS,CXJS,UNIT,IOP1,NAME,PH,IOP1,U,DHAT,F,FO,IFIRST,FNAME,N,ICASE,IOP1,IOP12,NOBS,F,INC,PRINT,NO)

WRITER FOR ONE SAMPLE ANALYSIS

- FUNCTION FOC-MC(X)
FUNCTION FOC-SP(X)
FUNCTION FOC-VAL(X)
FUNCTION FOC-VEP(X)
FUNCTION FOC-VG(X)
FUNCTION FOC-VG1(X)
FUNCTION FOC-VG2(X)
FUNCTION FOC-VG3(X)

UTILITY SUBROUTINES

SUBROUTINE ACDEF(INTAPE,NAME,IFORM,X)

SUBROUTINE TO CONVERT 8 CHARACTER ALPHANUMERIC ARRAY NAME WHICH IS IN A-FORMAT TO THE REAL VARIABLE X WHICH HAS THE 8 CHARACTER F-FORMAT IFOBT

SUBROUTINE FCODE(INTAPE,X,IFORM,NAME)

SUBROUTINE TO CONVERT REAL VARIABLE X WHICH HAS 4 CHARACTER F-FORMAT IFORM TO 8 CHARACTER ALPHANUMERIC ARRAY NAME WHICH IS IN A-FORMAT.

SUBROUTINE ICODE(INTAPE,K,IFORM,NAME)

SUBROUTINE TO CONVERT INTEGER VARIABLE K WHICH HAS 8 CHARACTER I-FORMAT IFORM TO 8 CHARACTER ALPHANUMERIC ARRAY NAME WHICH IS IN A-FORMAT.

SUBROUTINE JPLOT(X,Y,N,MM,NAMEX,NAMEY,NAME,ITITLE,IUNIT,IS3,IOP1, JSTART,I,I2,IOP1,IHEAD)

SUBROUTINE TO PRINT AND PRINTER PLOT THE VECTOR X, LISTING J AND OPTIONALLY A VECTOR Y, WHERE J IS THE SEQUENCE. MAX 50 (EVENLY DISTRIBUTED) VALUES ARE PLOTTED.

SUBROUTINES NOT ON THE ONESAM LIBRARY WHICH ARE USED IN THE ONESAM PACKAGE:

ADMRTS : IMSL SUBROUTINE TO COMPUTE INVERSE NORMAL DISTRIBUTION FUNCTION

DSHMAX : IMSL SUBROUTINE TO COMPUTE MIN AND MAX OF A VECTOR

DLPLT1 : TSOB SUBROUTINE TO PRINT AND PLOT A VECTOR (SEE JOE NEWTON FOR DOCUMENTATION)


```

25 CONTINUE
   IF(DOH(5)).EQ.0) GO TO 26
   IFIRST = 0
26 CONTINUE
   IF(DOH(6)).EQ.0) GO TO 27
   IF(PLOT.EQ.0)
     +CALL ONSAM(X,N,XO,ON,XS,UYS,CUXS,4,2,NAME,WIDTH,1,U,DHAT,F,FO,
     +FIRST,F0CAUC,ORDER,AMCAUCHY,IOP11,IOP12,NORS,FINC,XINT,NO)
   IFIRST = 0
27 CONTINUE
   IF(DOH(7)).EQ.0) GO TO 28
   IFIRST = 0
28 CONTINUE
   IF(DOH(8)).EQ.0) GO TO 29
   IF(PLOT.EQ.0)
     +CALL ONSAM(X,N,XO,ON,XS,UYS,CUXS,4,2,NAME,WIDTH,1,U,DHAT,F,FO,
     +FIRST,F0LGN0,ORDER,8HLOG NORM,IOP11,IOP12,NORS,FINC,XINT,NO)
   IFIRST = 0
29 CONTINUE
   IF(DOH(9)).EQ.0) GO TO 30
   IF(BETAP.GT.0.) GO TO 101
   WRITE(6,902) BETAP
101 CONTINUE
   IF(PLOT.EQ.0)
     +CALL ONSAM(X,N,XO,ON,XS,UYS,CUXS,4,2,NAME,WIDTH,1,U,DHAT,F,FO,
     +FIRST,F0PARE,ORDER,dIPARETO,IOP11,IOP12,NORS,FINC,XINT,NO)
   IFIRST = 0
30 CONTINUE
   IF(DOH(10)).EQ.0) GO TO 31
   IF(SETAU.GT.0.) GO TO 102
   WRITE(6,902) BETAU
102 CONTINUE
   IF(PLOT.EQ.0)
     +CALL ONSAM(X,N,XO,ON,XS,UYS,CUXS,4,2,NAME,WIDTH,1,U,DHAT,F,FO,
     +FIRST,F0EIB,ORDER,7HUEIBULL,IOP11,IOP12,NORS,FINC,XINT,NO)
   IFIRST = 0
31 CONTINUE
   IF(DOH(11)).EQ.0) GO TO 32
   IF(PLOT.EQ.0)
     +CALL ONSAM(X,N,XO,ON,XS,UYS,CUXS,4,2,NAME,WIDTH,1,U,DHAT,F,FO,
     +FIRST,F0MLGG,ORDER,BHIALF LBG,IOP11,IOP12,NORS,FINC,XINT,NO)
   IFIRST = 0
32 CONTINUE
999 CONTINUE
   STOP
901 FORMAT(/,5X,'SAMPLE SIZE OF ',15,' G.T. MAX OF 511. RUN ABORTED')
902 FORMAT(/,5X,F10.4,' IS ILLEGAL VALUE FOR BETA. DATA SET SKIPPED')
903 FORMAT(20F4.0)
   END

```

```

*****
FUNCTION ANEST(X,L,OPTKHM,OPTCOE)
*****
C
C *****
C FUNCTION TO COMPUTE AUTOREGRESSIVE ESTIMATOR EVALUATED AT X
C
C METHOD: ARTHO = OPTKHM / ABS(1 + Y)**2
C WHERE Y = OPTCOE(J)*EXP(I*J*2*PI*I*X) SUMMED OVER J = 1, L
C
C INPUT: X: SCALAR AT WHICH AUTOREGRESSIVE ESTIMATE IS EVALUATED.
C L: ORDER. MUST BE LESS THAN 11. SEE METHOD.
C OPTKHM: SEE METHOD.
C OPTCOE: AUTOREGRESSIVE COEFFICIENTS OF ORDER L. SEE METHOD.
C OPTCOE IS A COMPLEX 10-VECTOR.
C
C OUTPUT: FUNCTION RETURNS VALUE OF AUTOREGRESSIVE ESTIMATOR EVALUATED
C AT X.
C
C SUBROUTINES CALLED: NONE.
C *****
C
C *****
C COMPLEX OPTCOE(1)
C COMPLEX G
C PI=4.*ATAN(1.0)
C B=CNVALX(1.,0.)
C DO 1 J=1,L
C FJ=J
C G=G+OPTCOE(J)*DEXP(CMPLX(0.,X*2.*PI*I*FJ))
1 CONTINUE
C ANEST=OPTKHM/REAL(G+CONJ(G))
C RETURN
C END

```

```

SUBROUTINE AUTORG(A,LS,M,ALPHA,PHI,FKH)
C.....
C COMPUTES THE COEFFICIENTS ALPHA(L) AND FKH OF THE
C AUTOREGRESSIVE ESTIMATOR ACCORDING TO A RECURSIVE
C ALGORITHM
C INPUT :
C A : VECTOR OF COMPLEX FOURIER TRANSFORM,
C OF DIMENSION AT LEAST M
C M : (M-1) IS THE MAXIMUM ORDER OF SCHEME
C TO BE COMPUTED
C LS : ORDER OF SCHEME BEING COMPUTED. LS.GE.1
C OUTPUT :
C ALPHA : VECTOR OF COEFFICIENTS DEFINING THE
C APPROXIMATE FUNCTION, HAS TO BE DIMEN-
C SIONED AT LEAST M AND DECLARED COMPLEX
C FKH : SCALES THE AUTOREGRESSIVE ESTIMATOR TO
C INTEGRATE TO 1, DECLARED REAL
C ALPHA, PHI AND FKH ARE USED RECURSIVELY, I.E. THEIR
C VALUES AT OUTPUT FOR ORDER J ARE USED AS INPUT
C FOR ORDER (J+1)
C.....

```

```

C.....
C COMPLEX A(1),ALPHA(1),PHI(1),G,FJH
C COMPLEX FKH
C TROPI=8.*ATAN(1.0)
C FJH=CMPLX(0.,0.)
C PHI(1)=CMPLX(1.,0.)
C IF(LS.EQ.1) FKH=CONJG(A(1))
C DO 4 I=1,LS
C 4 FJH=FJH+CONJG(A(I+1))*PHI(I)
C G=FJH/FKH
C ALPHA(LS)=G
C IF(LS.EQ.1) GO TO 5
C K=LS-1
C DO 2 I=1,K
C 2 ALPHA(I)=ALPHA(I)+G*PHI(I)
C 5 CONTINUE
C 3 PHI(I)=CONJG(ALPHA(LS+1-I))
C FKH=FJH+CONJG(FKH)/CONJG(FKH)
C RETURN
C END

```

```

SUBROUTINE DATSTD(X,XO,M,ON)
C.....
C SUBROUTINE TO STANDARDIZE VECTOR X BY (X(I)-XO)/(X(N)-XO)
C AND RETURN THE STANDARDIZED DATA IN THE VECTOR ON
C INPUT :
C X,N : VECTOR OF ORDERED DATA OF LENGTH N
C XO : MINIMUM VALUE TO BE USED IN STANDARDIZATION
C OUTPUT :
C ON : VECTOR OF STANDARDIZED DATA OF LENGTH N+1
C SUBROUTINES CALLED : NCR/E
C.....
C DIMENSION X(N),ON(N)
C FACT = 1.0/(X(N) - XO)
C ON(1) = 0.0
C DO 10 I = 1, N
C 10 ON(I+1) = (X(I) - XO)*FACT
C RETURN
C END

```

```

SUBROUTINE DESTAT(X,N,NAME,JUNIT,IHEAD)
C.....
C SUBROUTINE TO PRINT ORDERED ARRAY BY QUANTILES AND COMPUTE
C DESCRIPTIVE STATISTICS.
C INPUT :
C X: ARRAY OF ORDER STATISTICS
C N: DIMENSION OF ARRAY X
C NAME: NAME OF DATA SET. MUST BE ARRAY OF DIMENSION 20 IN
C CALLING PROGRAM.
C JUNIT: NUMBER OF UNIT OUTPUT IS DESIRED ON.
C OUTPUT: PRINTED OUTPUT IS ON IUNIT.
C NO SUBROUTINES CALLED.
C.....
C DIMENSION X(N),NAME(20),SUM(4),SUMSQ(4)
C DIMENSION ALF(3)
C DIMENSION BOX(5),QUANT(5),ITITL(20),L(4)
C DIMENSION IHEAD(2),NAMEY(2),NAMEZ(2)
C DATA ITITL/HEAD,4,PILOT,18*4H /
C DATA NAMEY/ORDER,4H TITLE/
C DATA NAMEZ/4H 4H ON/
C DATA ALF/.05,.10,.25/

```

C COMPUTE AND PRINT DESCRIPTIVE STATISTICS.

```

K = 1
KK = 0
DO 50 I = 1,4
  KK = KK + L(I)
SUM(I) = 0.0
SUMSQ(I) = 0.0
DO 40 J = K, KK
  SUM(I) = SUM(I) + X(J)
  SUMSQ(I) = SUMSQ(I) + X(J)*X(J)
40 CONTINUE
K = K + L(I)
50 CONTINUE
WRITE(UNIT,1010) (SUM (I), I=1,4)
WRITE(UNIT,1011) (SUMSQ (I), I=1,4)
S = SUM(1) + SUM(2) + SUM(3) + SUM(4)
XBAR = S/FL0AT(N)
XB14 = (SUM(2) + SUM(3))/FL0AT(L2 + L3)
XB04 = (SUM(1) + SUM(4))/FL0AT(L1 + L4)
SS0 = SUMSQ(1) + SUMSQ(2) + SUMSQ(3) + SUMSQ(4)
VAR = (SS0 - S*S/FL0AT(N))/FL0AT(N-1)
SD = SORT(VAR)
R = X(L1+L2+L3+1) - X(L1)
XMED = (X(N/2) + MOD(N,2)) * X(N/2+1) /2.
WRITE(UNIT,1014) N,XMED
TRIM = (X(L1) + 2.*XMED + X(L1+L2+L3+1))/4.
GASTY = .3*X(N/3+1) + .4*XMED + .3*X(N-N/3)
WRITE(UNIT,1019) TRIM
WRITE(UNIT,1016) GASTY
DO 60 I = 1,3
  IG = INT(ALF(I)*FL0AT(N))
  N = N - 2*IG
  WGR1 = N - IG - 1
  IGR2 = IG + 2
  TRM = X(IG+1)*X(N-IG)
  WH = X(IG+1)+FL0AT(IG) + X(N-IG)*FL0AT(IG)
DO 70 J = IGR2,HWGR1
  TRM = TRM + X(J)
  WH = WH + X(J)
70 CONTINUE
TRM = TRM/H
WH = WH/FL0AT(N)
WRITE(UNIT,1017) ALF(I),WH
WRITE(UNIT,1018) ALF(I),TRM
60 CONTINUE
BOX(1) = X(I)
BOX(2) = X(L1)
BOX(3) = XIEB
BOX(4) = X(L1 + L2 + L3 + 1)
BOX(5) = X(N)
DO 90 I = 1,5
  QUANT(I) = FL0AT(I-1)*25.
CALL JPL0T(BOX,QUANT,5,80,NAMEB,NAMEY,NAMEZ,ITITL,UNIT,1,2,1,
  BOX,1H ,1,1,ICAD)
90

```

C COMPUTE L, THE ARRAY OF QUANTILE SIZES

```

L = N/4
LERR = N - 4*LL + 1
IF(LERR.EQ.1)GO TO 10
IF(LERR.EQ.2)GO TO 11
IF(LERR.EQ.3)GO TO 12
IF(LERR.EQ.4)GO TO 13
GO TO 999
10 CONTINUE
L1 = LL
L2 = LL
L3 = LL
L4 = LL
GO TO 20
11 CONTINUE
L1 = LL
L2 = LL
L3 = LL
L4 = LL + 1
GO TO 20
12 CONTINUE
L1 = LL + 1
L2 = LL
L3 = LL
L4 = LL + 1
GO TO 20
13 CONTINUE
L1 = LL
L2 = LL + 1
L3 = LL + 1
L4 = LL + 1
GO TO 20
20 CONTINUE
L(1)=L1
L(2)=L2
L(3)=L3
L(4)=L4
C
C PRINT DATA ARRAY - ONE COLUMN FOR EACH QUANTILE.
WRITE(UNIT,1001) NAME
WRITE(UNIT,1020) IHEAD
WRITE(UNIT,1002)
WRITE(UNIT,1003)
DO 30 I = 1,LL
  X(I) = X(L1+L2+1),
  X(L1+L2+L3+1)
  WRITE(UNIT,1005)
  IF(L1.GT.LL) WRITE(UNIT,1006) X(L1)
  IF(L2.GT.LL) WRITE(UNIT,1007) X(L2 + L1)
  IF(L3.GT.LL) WRITE(UNIT,1008) X(L3 + L2 + L1)
  IF(L4.GT.LL) WRITE(UNIT,1009) X(L4 + L3 + L2 + L1)
  IF(L4.GT.LL) WRITE(UNIT,1015) L4
30

```

999 CONTINUE
RETURN

```

C
1001 FORMAT(IN, 'DATA SET: ',20A4)
1002 FORMAT(INO,20X, 'ORDER STATISTICS IN QUARTERS',/1H*,20X,
      * ' FIRST QUARTER SECOND QUARTER THIRD QUARTER FOURTH QUARTER' /
      * '-----' /
      * ' ' /
      * ' ' /
      * ' ' /
1004 FORMAT(IX,10 , 4(IX,F15.4))
1005 FORMAT(IX)
1006 FORMAT(IX, 2X,F15.4)
1007 FORMAT(IN, 2X,F15.4)
1008 FORMAT(IN, 4IX,F15.4)
1009 FORMAT(IN, 2X,F15.4)
1010 FORMAT(IX, 'SUM', 3X,4(IX,F15.4))
1011 FORMAT(IX, 'S.D.', 3X,4(IX,F15.4))
1012 FORMAT(IX, 'SUM Q', 3X,4(IX,F15.4))
1013 FORMAT(IX, 'MEAN', 3X,4(IX,F15.4))
1014 FORMAT(IX, 'VARIANCE', 3X,4(IX,F15.4))
1015 FORMAT(IX, 'STANDARD DEVIATION', 3X,4(IX,F15.4))
1016 FORMAT(IX, 'INTERQUARTILE RANGE', 3X,4(IX,F15.4))
1017 FORMAT(IX, 'TRIMMED MEAN', 3X,4(IX,F15.4))
1018 FORMAT(IX, 'WINGED MEAN', 3X,4(IX,F15.4))
1019 FORMAT(IX, 'TRINKLED MEAN', 3X,4(IX,F15.4))
1020 FORMAT(IX, 'END')
END

```

```

C *****
C SUBROUTINE FOURIER(F, N, M, A, MA)
C *****
C SUBROUTINE TO COMPUTE THE FOURIER TRANSFORM
C PHI(V) OF A DENSITY DEFINED ON (0,1) FOR V=0,1,...,M
C INPUT :
C F,U,M : VECTORS OF LENGTH N CONTAINING F(U),U
C MA : MAXIMUM VALUE OF V FOR WHICH PHI(V) IS COMPUTED
C OUTPUT : A : COMPLEX-VALUED VECTOR CONTAINING THE PHI'S
C SUBROUTINES CALLED : NONE
C *****
C DIMENSION F(1),U(1)
C COMPLEX A(MA),Z
C THOPI=8.*ATAN(1.)
C FN=FLOAT(N)
C NO 20 IM=1,MA
C FIR=IM-1
C A(IM)=CMPLX(0.,0.)
C DO 10 I=1,M
C Z=CMPLX(0.,TWOPI*F*IM*U(I))
10 A(IM)=A(IM)+F(I)*EXP(Z)
C 20 CONTINUE
C A(1)=CMPLX(1.,0.)
C RETURN
C END

```



```

SUBROUTINE PARZ(RVAR,M,N,CAT,NORD)
C
C SUBROUTINE TO DETERMINE THE ORDER OF AN AUTOREGRESSIVE
C PROCESS BY PARZEN'S CRITERIA
C
C INPUT :
C M,RVAR(1),...,RVAR(N) : STANDARDIZED RES VAR
C FOR ORDERS 1 THRU M.
C N : SAMPLE SIZE
C
C OUTPUT :
C NORD : DETERMINED ORDER
C CAT(1),...,CAT(N)
C
C SUBROUTINES CALLED : MIN
C
C DIMENSION RVAR(N),CAT(N)
C
C ON=FLOAT(N)
C DO 1 I=1,M
C O=0.
C DO 2 J=1,I
C O=C*(1.-(FLOAT(J)/ON))/RVAR(J)
C O=C/O
C
C 1 CAT(I)=C-(1.-(FLOAT(I)/ON))/RVAR(I)
C CALL MINCAT,M,CAT(N),NORD
C IF(CAT(N).GT.-1.) NORD=0
C
C RETURN
C END
C
SUBROUTINE PRINTA(A,N,IUNIT)
C
C SUBROUTINE TO COMPUTE AND PRINT THE SQUARE MODULUS OF THE
C COMPLEX-VALUED FOURIER TRANSFORMS A(1),...,A(N)
C
C INPUT :
C A,N
C IUNIT : UNIT ON WHICH OUTPUT IS DESIRED :
C
C SUBROUTINES CALLED : NONE
C
C DIMENSION A(N)
C COMPLEX A(N)
C
C WRITE(IUNIT,1000)
C DO 1 I=1,N
C 10 A(I) = REVL(COJUS(A(I))HA(I))
C CALL COLFCT(A,N,I,NPHNIZ)
C RETURN
C
1000 F=FLOAT(///5X,'MODULUS SQUARED OF FOURIER COEFFICIENTS'///)
C END

```

```

SUBROUTINE OMOX(X,Y,NM,NAME)
C
C SUBROUTINE OBTAINS PRINTER PLOT OF QUANTILE BOX PLOT WITH
C MEDIAN LABELED, AND LINES AT RANGES, EIGHTHS, 1/40 1/8THS
C
C INPUT: X,NM : DATA VECTOR OF LENGTH NM=2*N+1 (IE EXPANDED
C DATA SET)
C Y : WORK VECTOR OF LENGTH NM
C NAME : VECTOR CONTAINING NAME OF DATA SET
C
C SUBROUTINES CALLED : NONE
C
C DIMENSION X(NM),Y(NM),INDU(80),KP(7),NAME(20)
C DATA IB,IM,IX,IB,12,13,14,15/1H,16H,17H,18H,19H,20H,11H,12H,13H,14H,15H,16H,17H,18H,19H,20H
C 1HS,11H,12H,13H/
C NI=NM-1
C XI=X(NI)-X(1)
C IX=X(NI)/XI
C DO 10 J=1,NM
C Y(J)=X(J)
C X(J)=(X(J)-X(NI))/IX
C M=X(J)*79.
C X(J)=80-M
C 10 CONTINUE
C
C KP(1)=.0625*FLOAT(NM)+1.
C KP(2)=.125*FLOAT(NM)+1.
C KP(3)=.25*FLOAT(NM)+1.
C KP(4)=.5*FLOAT(NM)+1.
C KP(5)=NM-KP(3)+2
C KP(6)=NM-KP(2)+2
C KP(7)=NM-KP(1)+2
C
C WRITE(6,1000)NAME
C J=1
C I=1
C 12 CONTINUE
C DO 13 K=1,80
C 13 INDU(K)=I
C 15 CONTINUE
C J=NM-I+1
C IF(X(J).NE.FLOAT(I)/NM) GO TO 50
C K=(I+NM-I)/NM+79
C K=80-K
C DO 20 L=1,7
C 20 IF(JJ.EQ.KP(L)) GO TO 100
C 40 CONTINUE
C IF(INDU(K).EQ.INDU TO 45
C IF(INDU(K).EQ.INDU TO 140
C IF(INDU(K).EQ.INDU TO 142
C IF(INDU(K).EQ.INDU TO 144
C IF(INDU(K).EQ.INDU TO 146
C IF(INDU(K).EQ.INDU TO 148

```

```

SUBROUTINE DMPLOT(X,MM1,IHEAD,IOP1,J,IBATCH,LAB)
C.....
C SUBROUTINE OBTAINS CALCOMP PLOT OF QUANTILE BOX PLOT, WITH
C IOP1 BOXES DRAWN AND A CONFIDENCE INTERVAL FOR THE MEDIAN DRAWN
C INPUT: X,MM1 : DATA VECTOR OF LENGTH MM1-2+MM1 (IE EXPANDED
C DATA SET)
C IHEAD : VECTOR CONTAINS NAME OF DATA SET, USED AS LABEL
C IOP1 : = NUMBER OF BOXES DESIRED (CALC AT H,L,D,ETC.
C IBATCH : = NUMBER OF BATCHES ANALYZED. IBATCH+1 GIVES
C A SQUARE BOX PLOT 10X10. IBATCH GT 1 GIVES BOX PLOT AX10
C U : WORK VECTOR OF LENGTH MM1
C LAB : CONTAINS SUBHEADING OF LE 4 CHARACTERS
C SUBROUTINES CALLED : MMNSC
C CALCOMP SUBROUTINES CALLED : PLOT, AXIS1, LINE, SYMBOL
C.....
DIMENSION X(1),U(1),IHEAD(20),Z(7),UZ(7)
MM1=MM1-1
MM1H=MM1-1
MM1L=X(1)
MM1R=X(MM1)
DO 5 I=1,MM1H
  U(I)=FLOAT(I)/FLOAT(MM1)
  X(I)=X(I+1)
5 CONTINUE
  X(1)=X(2)
  U(1)=U(2)
  ZLE=10.
  ZINC=.1
  IF (IBATCH.GT.1) ZLEN=4.0
  IF (IBATCH.GT.1) ZINC=.25
  CALL MMNSC(XMIN,XMAX,10.,A,B)
  CALL AXIS1(0.,0.,U,-1,XLEN,0.,0.,XINC)
  CALL AXIS1(0.,0.,U(8),4,10.,96.,8,B)
  X(MM1)=A
  X(MM1)=B
  U(MM1)=0.
  U(MM1)=XINC
  CALL LINE(U,X,MM1H,1,2,4)
  UZ(1)=0.
  UZ(2)=1.
  UZ(3)=0.
  UZ(4)=XINC
  Z(1)=XMIN
  Z(2)=XMAX
  Z(3)=A
  Z(4)=B
  CALL LINE(UZ,Z,2,1,-1,4)
  CALL SYMBOL(.2,10.2,.175,LAB,0.,10)
  CALL SYMBOL(.2,10.5,.28,LINEAB,0.,88)
  DO 50 J=1,IOP1
  K1=FLOAT(J)/Z.(Z(J+1)-Z(J))
  K2=MM1-K1

```

```

INDUCK=IX
60 TO 45
140 INDUCK=12
60 TO 45
142 INDUCK=13
60 TO 45
144 INDUCK=14
60 TO 45
146 INDUCK=15
60 TO 45
45 CONTINUE
  J=J+1
  IF (J.GT.MM1) GO TO 50
50 CONTINUE
  U=FLOAT(J)/FLOAT(MM1)
  WRITE(6,1005)U,Y(J+1),MMU
  I=I+1
  IF (I.GT.50) GO TO 40
  IF (J.GT.MM1) GO TO 40
60 TO 12
140 CONTINUE
  K1=(LEQ.4)GO TO 110
  K1=(FLOAT(MM1-U-1))/FLOAT(MM1)*79
  K1=80-K1
  K2=MM1-K1
  K1R=K1/200(K1,K)
  K2R=K2/200(K1,K)
  K1R=K1-1
  DO 100 I=K1R,K2R,MMAX
  INDUCK=I
  IF (INDUCK=1).EQ.10 INDUCK(MM1-1)=I
  IF (INDUCK(MM1).EQ.1) INDUCK(MMAX+1)=I
60 TO 40
110 INDUCK=I+1
60 TO 45
40 CONTINUE
  WRITE(6,1008)
  DO 50 J=1,IOP1
  60 X(J)=Y(J)
500 FORMAT(1X,1X,DATA SET,/,20A4,/,1X,QUANTILE-BOX PLOT,/,
  4X,U,5X,UB),/,1X,14(IH-),1X,LEQ(IH-))
1005 FORMAT(1X,5X,J,F10.3,1X,DATA SET,/,
  1008 FORMAT(1X,15(IH-),1X,60(IH-))
C
RETURN
END

```

```

UZ(1)=FLOAT(K1)/FLOAT(NH)
UZ(2)=FLOAT(K2)/FLOAT(NH)
UZ(3)=UZ(2)
UZ(4)=UZ(1)
UZ(5)=UZ(1)
UZ(6)=0.
UZ(7)=XINC
Z(1)=X(61)
Z(2)=Z(1)
Z(3)=X(62)
Z(4)=Z(3)
Z(5)=Z(1)
Z(6)=A
Z(7)=B
CALL LINE(UZ,Z,5,1,0,0)
30 CONTINUE
40 CONTINUE
50=FLOAT(NH)*.25
U=(FLOAT(NH)/FLOAT(NH))*XLEN
UD=(FLOAT(NH)/FLOAT(NH))*XLEN
EM=FLOAT(NH)*.5
X=(X(1N)-A)/B
US=(FLOAT(NH)/FLOAT(NH))*XLEN
CALL PLOT(U, A, 3)
CALL PLOT(U, A, 2)
X=(X(6N)-X(61))/SORT(FLOAT(NH))/B
XNN=N-N
IF(XNN.LT.0.)XNN=0.
XPP=X*DXH
IF(XPP.GT.10.)XPP=10.
CALL PLOT(U, XPP, 3)
CALL PLOT(U, XPP, 2)
CALL PLOT(-1., -1., -3)
DO 70 I=1, NEN
  IN=IN-1
  X(I)=X(IN)
  X(NH)=X(NH)
  XLEN=XLEN+1.
CALL PLOT(X(I), 0., -3)
RETURN
END

```

```

SUBROUTINE QOING(X,N,IOP1,NAME)
C.....
C SUBROUTINE ORDERS, PRINTS DATA, EXPANDS DATA SET BY COMPUTING
C AVERAGES OF CONSECUTIVE DATA PIS AND COMPUTES DIMONISTICS
C INPUT: X,N : DATA VECTOR OF LENGTH N. MUST BE DIMENSIONED
C AT LEAST 2*N+1. ON OUTPUT, X CONTAINS THE EXPANDED
C DATA SET
C IOP1 : =1 IF USER WISHES TO INPUT A NATURAL PER IN
C X(N+1) AND NATURAL MAX IN X(N+2), EG WHEN USING
C SEVERAL BATCHES
C NAME : VECTOR CONTAINING NAME OF DATA SET USED FOR LABEL
C SUBROUTINES CALLED : QUICK, XBARS2(ITSBD), HONKIS(IMS.)
C.....
DIMENSION X(1),NAME(20)
WRITE(6,1001)NAME,N
WRITE(6,1002)(X(I),I=1,N)
CALL XBAR(S2(X,N,XBAR,S2)
SD=SQRT(S2)
IF(IOP1.EQ.1)XNIN=X(N+1)
IF(IOP1.EQ.1)XMAX=X(N+2)
CALL QUICK(N,X)
WRITE(6,1009)
WRITE(6,1002)(X(I),I=1,N)
NN=2*N
N1=NN+1
N1=N-1
DO 10 J=1,N
  J1=J-1
  JN=2*(N-J1)
  JN=2*N-2*J1-1
  NJ=N-J1
  NJ1=N-J1-1
  IF(NJ1.EQ.0)NJ1=1
  X(JNE)=X(NJ)
  X(JND)=(X(NJ)+X(NJ1))/2.
10 CONTINUE
  X(NR1)=X(NR)
  IF(IOP1.EQ.1)X(1)=X(NIN)
  IF(IOP1.EQ.1)X(NH)=XMAX
  K1=-.0625*FLOAT(NH)
  K2=-.125*FLOAT(NH)
  K3=-.25*FLOAT(NH)
  K4=-.5*FLOAT(NH)
  K5=N-K3
  K6=N-K2
  K7=N-K1
  WRITE(6,1008)
  WRITE(6,1003)K7,X(K7+1),K6,X(K6+1),K5,X(K5+1),K4,X(K4+1),K3,
  *X(K3+1),K2,X(K2+1),K1,X(K1+1)
  NENID=(X(K5+1)+X(K3+1))/2.
  ENID=(X(K6+1)+X(K2+1))/2.
  B3ID=(X(K7+1)+X(K1+1))/2.
  AVMID=(X(K5+1)+X(K3+1)+X(K6+1)+X(K2+1))/4.
  HHR=(X(K5+1)+X(K3+1))/SORT(FLOAT(NH))

```

```

WRITE(6,1004)X(K(4:1)),MMID,EMID,DMID,AVRID,XBAR,JHIN
CALL MDARIS(.75,LENN,HEX)
XEND=(X(US*1)-X(K(3:1)))/I2.*DENN)
CALL MDARIS(.875,GELE,IER)
XEND=(X(K(6:1))-X(K(2:1)))/I2.*DEME)
CALL MDARIS(.9375,REB,IER)
XEND=(X(K(7:1))-X(K(1:1)))/I2.*DEBD)
AEND=(XEND+MEFOR+XENOR)/3.
XESP=(X(US*1)-X(K(3:1)))/ALOG(3.)
XEEP=(X(US*1)-X(K(2:1)))/ALOG(7.)
XEXP=(X(K(7:1))-X(K(1:1)))/ALOG(15.)
AEXP=(XEXP+MEFOR+XENOR)/3.
WRITE(6,1005)XBAR,XEYP,XENOR,XEYP,XENOR,XEYP,XENOR,XEYP,XENOR,XEYP,SD
XEN=(X(K(4:1))-X(K(3:1)))/X(K(3:1))
XEP=(X(US*1)-X(K(2:1)))/X(K(2:1))
XEN=(X(K(6:1))-X(K(5:1)))/X(K(5:1))
XEP=(X(K(7:1))-X(K(6:1)))/X(K(6:1))
WRITE(6,1006)XEN,XEP,XEN,XEP
XEN=ALOG(X(K(6:1))-X(K(5:1)))/X(K(5:1))
XEP=ALOG(X(K(7:1))-X(K(6:1)))/X(K(6:1))
XEN=XEN/LOG(XEN/MEFOR)
XEP=XEP/LOG(XEP/MEFOR)
XEN=ALOG(XEN/LOG(XEN/MEFOR))
XEP=ALOG(XEP/LOG(XEP/MEFOR))
WRITE(6,1007)XEN,XEYP,XENOR,XEYP,XENOR,XEYP,XENOR,XEYP,XENOR,XEYP,SD
RETURN
1001 FORMAT(1H1,/,IX,DATA SET: ',20A4,/,IX,SAMPLE SIZE = ',I5,/)
1002 FORMAT(3F12.3)
1003 FORMAT(7,/,IX,50(1H9),/,IX,7 POINT SUMMARY',/
+3X,UPPER D = X(',I4,') = ',F12.3,/,
+3X,UPPER E = X(',I4,') = ',F12.3,/,
+3X,UPPER HINGE = X(',I4,') = ',F12.3,/,
+3X,MEDIAN = X(',I4,') = ',F12.3,/,
+3X,LOWER HINGE = X(',I4,') = ',F12.3,/,
+3X,LOWER E = X(',I4,') = ',F12.3,/,
+3X,LOWER D = X(',I4,') = ',F12.3,/)
1004 FORMAT(1X,MISUMMARIES',/,2X,MEDIAN = ',F12.3,/,
+2X,MID = ',F12.3,/,2X,AVE = ',F12.3,/,
+2X,XBAR = ',F12.3,/,2X,AVE = ',F12.3,/,
+2X,XBAR = ',F12.3,/,2X,MISORTIN) = ',F12.3,/)
1005 FORMAT(1X,SCALE SUMMARIES',/
+2X,MINUSNORMAL = ',F12.3,/,2X,MINUSNORMAL = ',F12.3,/,
+2X,SAMPLE ST DEV = ',F12.3,/)
1006 FORMAT(1X,SUMMARY STATISTICS',/,2X,MINUSNORMAL = ',F7.4,/,
+2X,MINUSNORMAL = ',F7.4,/,2X,MINUSNORMAL = ',F7.4,/)
1007 FORMAT(1X,TAIL DIAGNOSTIC MEASURES',/,2X,SAMPLE 'SX,NORMAL(0,1)
+3X,'EXP(1)',/,2X,LOG(CHEB) = ',F12.4,/,2X,LOG(CHE/BD) = ',
+3F12.4,/,/,IX,50(1H9),/)
1008 FORMAT(1H1)
1009 FORMAT(7,/,IX,ORDERED DATA',/)
END

```

```

SUBROUTINE GRIN(X,N,M,U,Q,IOP1,NO,MMIN,XINT)
C *****
C SUBROUTINE TO COMPUTE Q(U) AT U = H,2H,3H,....,H WHERE
C H IS > OR = .002 BY LINEAR INTERPOLATION FROM OTHER
C STATISTICS IF DATA IS UNGROUPED (IOP1=0), OR FROM TALLIES
C IF DATA IS GROUPED (IOP1=1).
C INPUT : X : VECTOR OF ORDER STATISTICS IF DATA IS UNGROUPED
C          VECTOR OF TALLIES IF DATA IS GROUPED
C          N : DIMENSION OF X
C          XMIN : NATURAL MINIMUM (IF DATA IS UNGROUPED)
C                LOWER LIMIT OF 1ST GROUP (IF DATA IS GROUPED)
C          XINT : GROUP WIDTH (IF DATA IS GROUPED)
C          IOP1 : = 0 IF DATA IS UNGROUPED
C                = 1 IF DATA IS GROUPED
C          H : FIRST VALUE (AND INCREMENT) FOR COMPUTING Q
C                USUALLY H=.01,.025, OR .05
C OUTPUT : U : VECTOR OF U VALUES (MAX 500)
C          Q : VECTOR CONTAINING Q(U) (MAX 500)
C SUBROUTINES CALLED : NONE
C *****
C DIMENSION X(1),U(1),Q(1)
C NO=1./H
C IF (NO.GT.500)GO TO 900
C IF ((FLOAT(NO)+H).EQ.1.) NO=NO-1
C IF (IOP1.EQ.1)GO TO 50
C COMPUTE Q FOR UNGROUPED DATA
C X(N+1)=X(N)
C DO 10 I=1,NO
C J=FLOAT(N)+FLOAT(I)+H*.5
C U(I)=FLOAT(I)+H
C A=U(I)+FLOAT(N)-FLOAT(J)+.5
C Q(I)=A+X(J)+((1.-A)*X(J))
C 10 CONTINUE
C GO TO 80
C 50 CONTINUE
C COMPUTE Q FOR GROUPED DATA
C XDE=0.
C DO 60 I=1,M
C XDE=XDE+X(I)
C 60 CONTINUE
C U=0.
C XUM=0.
C IU=1
C IV=1
C UC(1)=H
C 65 XUM=XUM+X(IU)
C U=UUM/XUM

```

```

70 IF (U(1), U(2), U(3), U(4), U(5), U(6), U(7), U(8), U(9), U(10)) GO TO 75
   U=U+1
   U=U
   GO TO 65
75 CONTINUE
   A=(U(10)-U(1))/(U(10)-U(1))
   X=X1+A*(X2-X1)
   U(10)=A*X1+(1-A)*X2
   U=U+1
   IF (U(10)-U(1)) GO TO 60
   U(10)=FLOOR(U(10)*N)
   GO TO 70
80 CONTINUE
   DO 90 I=1, N
   I2=NQ-I+2
   I1=NQ-I+1
   U(I2)=U(I1)
   U(I1)=X*MIN
   U(I1)=0.
   GO TO 99
900 WRITE(6,900)N
901 FORMAT(//5X, 'INCREMENT OF .F3.A. IS TOO SMALL TO COMPUTE',
//, '0.1' INCREMENT MUST BE AT LEAST .002')
999 CONTINUE
   RETURN
   END
SUBROUTINE QTOF(Q,U,NQ,XS,FO)
.....
C SUBROUTINE TO COMPUTE Q(U) AND FQ(U)=1./Q(U) FROM
C THE EMPIRICAL QUANTILE FN CAP Q(U) AND THE U VALUES.
C INPUT : Q,U,NQ
C OUTPUT : XS : SPACINGS-Q(U)
           FO
.....
C DIMENSION Q(1),U(1),XS(NQ),FQ(NQ)
   XS(1)=Q(NQ+1)-Q(NQ)/(U(NQ+1)-U(NQ))
   XS(NQ)=XS(NQ)
   FQ(1)=1.
   DO 10 I=1, NQ-1
   XS(I)=(Q(I+2)-Q(I))/(U(I+2)-U(I))
   IF (XS(I).EQ.0.) GO TO 10
   IF (XS(I).LT.XS(NQ)*MIN) XS(I)
10 CONTINUE
   XS(NQ)=XS(NQ)*MIN
   DO 20 I=1, NQ
   FQ(XS(I)+.001)=XS(I)-XS(NH)
   FQ(1)=1./Q(1)
20 CONTINUE
   RETURN
   END

```

```

SUBROUTINE QUICKIN,T)
.....
C QUICK SORT THIS ALGORITHM IS ALSO REFERRED TO AS A PARTICIATED
C EXCHANGE SORT. EXPECTED RUNTIME IS PROPORTIONAL TO N*LOG2(N)
C ALTHOUGH THE WORST CASE IS PROPORTIONAL TO N^2.
C REFERENCE: DONALD E. KNUTH- THE ART OF COMPUTER PROGRAMMING VOL 3.
C INPUT : X,N : VECTOR TO BE SORTED OF LENGTH N
C OUTPUT : X : SORTED VECTOR
C SUBROUTINES CALLED : NONE
.....
REAL T(N),Y
INTEGER IP,LV(16),IV(16),IP,IUP
LV(1)=1
IV(1)=N
IP=1
10 IF (IP.LT.1) GO TO 75
15 IF ((IV(IP)-LV(IP)).LT.1) GO TO 20
   GO TO 25
20 IP=IP-1
   GO TO 10
25 LP=LV(IP)-1
   IUP=IV(IP)
   Y=T(IUP)
   T(IUP)=T(LP)
   T(LP)=Y
30 IF ((IUP-LP).LT.2) GO TO 45
   LP=LP+1
   IF (T(LP).LE.Y) GO TO 30
   T(IUP)=T(LP)
35 IF ((IUP-LP).LT.2) GO TO 40
   IUP=IUP-1
   IF (T(IUP).GE.Y) GO TO 35
   T(LP)=T(IUP)
   GO TO 30
40 IUP=IUP-1
45 IUP=Y
   GO TO 60
55 LV(IP+1)=LV(IP)
   IV(IP+1)=IUP-1
   LV(IP)=IUP+1
   GO TO 70
60 LV(IP+1)=IUP+1
   IV(IP)=IUP-1
   GO TO 15
75 RETURN
   END

```

